

# Constraining the Astrophysics of the Cosmic Dawn and the Epoch of Reionization with the SKA

**Andrei Mesinger<sup>\*1</sup>, Andrea Ferrara<sup>1</sup>, Bradley Greig<sup>1</sup>, Ilian Iliev<sup>2</sup>, Garrelt Mellema<sup>3</sup>, Jonathan Pritchard<sup>4</sup>, Mario G. Santos<sup>5,6,7</sup>, on behalf of the EoR/CD Science Team**

<sup>1</sup>*Scuola Normale Superiore, Piazza dei Cavalieri, 7 56126 Pisa, Italy;* <sup>2</sup>*Astronomy Centre, Department of Physics and Astronomy, University of Sussex, Falmer, Brighton BN1 9QH, UK;*

<sup>3</sup>*Dept. of Astronomy & Oskar Klein Centre, Stockholm University, AlbaNova, SE-10691 Stockholm, Sweden;* <sup>4</sup>*Department of Physics, Blackett Laboratory, Imperial College, London SW7 2AZ, UK;* <sup>5</sup>*Department of Physics, University of Western Cape, Cape Town 7535, South Africa;* <sup>6</sup>*SKA SA, 4rd Floor, The Park, Park Road, Pinelands, 7405, South Africa;* <sup>7</sup>*CENTRA, Departamento de Fisica, Instituto Superior Tecnico, 1049-001 Lisboa, Portugal*

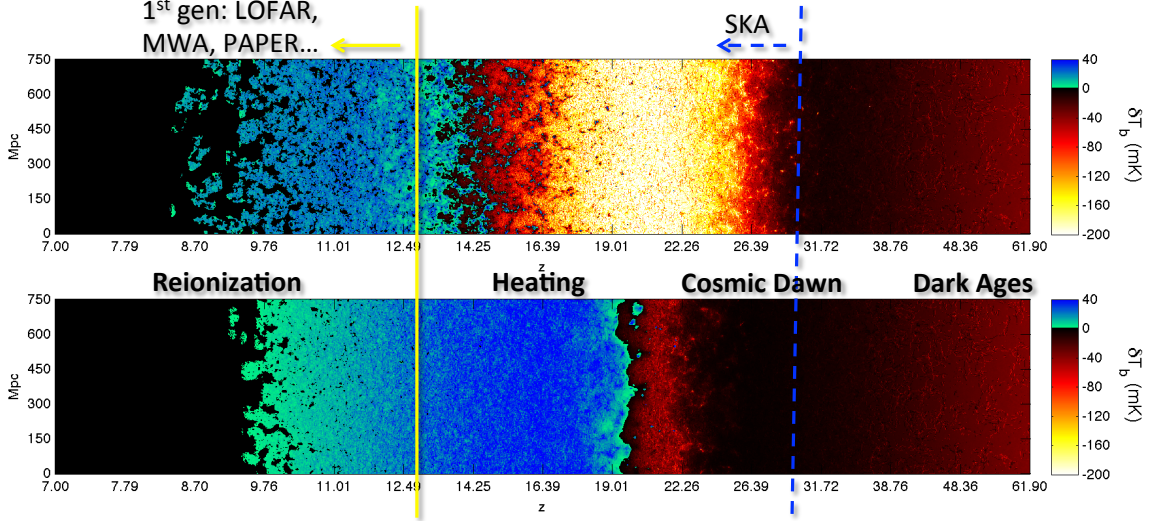
*E-mail: [andrei.mesinger@sns.it](mailto:andrei.mesinger@sns.it)*

The Square Kilometre Array (SKA) will offer an unprecedented view onto the early Universe, using interferometric observations of the redshifted 21cm line. The 21cm line probes the thermal and ionization state of the cosmic gas, which is governed by the birth and evolution of the first structures in our Universe. Here we show how the evolution of the 21cm signal will allow us to study when the first generations of galaxies appeared, what were their properties, and what was the structure of the intergalactic medium. We highlight qualitative trends which will offer robust insights into the early Universe.

*Advancing Astrophysics with the Square Kilometre Array  
June 8-13, 2014  
Giardini Naxos, Italy*

---

<sup>\*</sup>Speaker.



**Figure 1:** 21cm brightness temperature offset from the CMB (from [1]). The horizontal axis shows evolution along the comoving line-of-sight coordinate, from  $z \approx 62$  to  $z \approx 7$ . From right to left we see the expected major milestones in the signal: (i) collisional decoupling (*red*→*black*); (ii) WF coupling (*black*→*yellow*); (iii) IGM heating (*yellow*→*blue*); (iv) reionization (*blue*→*black*). The *top panel* corresponds to a “fiducial” model, while the *lower panel* corresponds to an “extreme X-ray” model in which primordial galaxies are much more efficient than local star-bursts in generating X-rays (see, e.g. [2]), saturating the unresolved soft X-ray background [3] by  $z \sim 10$ . The lower panel further assumes that soft X-rays with energies  $\lesssim 1\text{keV}$  are absorbed within the host galaxy. Although the models have comparable electron scattering optical depths,  $\tau_e$ , the astrophysical milestones are very different in 21cm. The redshift limit accessible to first generation interferometers roughly corresponds to the vertical yellow line; the SKA-LOW should probe out to roughly the vertical blue line, opening-up a new window on the Cosmic Dawn. The slices are 750 Mpc in height and 1.5 Mpc thick.

## 1. Introduction

With unprecedented resolution and sensitivity, the SKA-LOW will enable ground-breaking studies of early Universe astrophysics through the 21cm line from neutral hydrogen. No other planned instrument will allow us to study baryons at high redshift in such detail. We will be able to trace the delicate, physics-rich interplay between the intergalactic medium (IGM) and the first galaxies.

As a cosmological probe, the signal is usually represented in terms of the offset of the 21cm brightness temperature from the cosmic microwave background (CMB) temperature,  $T_\gamma$ , along a line of sight at observed frequency  $\nu$  (c.f. [4]):

$$\delta T_b(\nu) = \frac{T_S - T_\gamma}{1+z} (1 - e^{-\tau_{\nu_0}}) \approx \quad (1.1)$$

$$27 x_{\text{HI}} \left(1 - \frac{T_\gamma}{T_S}\right) (1 + \delta_{\text{nl}}) \left(\frac{H}{dv_r/dr + H}\right) \sqrt{\frac{1+z}{10}} \frac{0.15}{\Omega_M h^2} \left(\frac{\Omega_b h^2}{0.023}\right) \left(\frac{1-Y_p}{0.75}\right) \text{mK},$$

where  $T_S$  is the gas spin temperature,  $\tau_{\nu_0}$  is the optical depth at the 21cm frequency  $\nu_0$ ,  $\delta_{\text{nl}}(\mathbf{x}, z) \equiv \rho/\bar{\rho} - 1$  is the evolved (Eulerian) density contrast,  $Y_p$  is the Helium mass fraction,  $H(z)$  is the Hubble parameter,  $dv_r/dr$  is the comoving gradient of the line of sight component of the comoving velocity, and all quantities are evaluated at redshift  $z = \nu_0/\nu - 1$ . The cosmological 21cm signal uses the CMB as a back-light: if  $T_S < T_\gamma$ , then the gas is seen in absorption, while if  $T_S > T_\gamma$ , the gas is seen in emission.

The spin temperature,  $T_S$ , interpolates between the CMB temperature,  $T_\gamma$ , and the gas kinetic temperature,  $T_K$ . Since we only observe the *contrast* of the gas against the CMB, a signal is only obtained if  $T_S \rightarrow T_K$ . This coupling is achieved through either: (i) collisions, which are effective in the intergalactic medium (IGM) at high redshifts,  $z \gtrsim 50$ ; or (ii) a Lyman alpha background (so-called Wouthuysen-Field (WF) coupling; [5, 6]), effective soon after the first sources turn on.

In the top panel of Fig. 1, we show a slice through the  $\delta T_b$  field in a “fiducial” model (below we use the term “fiducial” to refer to models in which atomically-cooled galaxies have similar X-ray and UV properties as local ones, with an ionizing emissivity such that the mid-point of reionization is at  $z \sim 10$ ; for further details, see, e.g. [1]). It is immediately obvious that the 21cm signal is a physics-rich probe, encoding information on various processes during the Cosmic Dawn (CD) and Epoch of Reionization (EoR). Although the exact timing of the cosmic epochs is uncertain, the relative order is robustly predicted (c.f. [7]; §2.1 in [8]):

1. **Collisional coupling:** The IGM is dense at high redshifts, so the spin temperature is uniformly collisionally coupled to the gas kinetic temperature,  $T_K = T_S \lesssim T_\gamma$ . Following thermal decoupling from the CMB ( $z \lesssim 300$ ), the IGM cools adiabatically as  $T_K \propto (1+z)^2$ , faster than the CMB:  $T_\gamma \propto (1+z)$ . Thus  $\delta \bar{T}_b$  is negative. This epoch, serving as a *clean probe of the matter power spectrum* at  $z \gtrsim 100$ , is not shown in Fig. 1.
2. **Collisional decoupling:** The IGM becomes less dense as the Universe expands. The spin temperature starts to decouple from the kinetic temperature, and begins to approach the CMB temperature again,  $T_K < T_S \lesssim T_\gamma$ . Thus  $\delta \bar{T}_b$  starts rising towards zero. Decoupling from  $T_K$  occurs as a function of the local gas density, with underdense regions decoupling first. Fluctuations are sourced by the density field, and again *offer a direct probe of cosmology*. Eventually ( $z \sim 25$ ), all of the IGM is decoupled and there is little or no signal. This epoch corresponds to the red→black transition on the right edge of Fig. 1.
3. **WF coupling (i.e. Ly $\alpha$  pumping):** The first astrophysical sources turn on, and begin coupling  $T_S$  and  $T_K$ , this time through the Ly $\alpha$  background.  $\delta \bar{T}_b$  becomes more negative, reaching values as low as  $\delta \bar{T}_b \sim -100 - -200$  mK (depending on the offset of the WF coupling and X-ray heating epochs). Fluctuations are driven by the strength of the Ly $\alpha$  background. This epoch, offering a window on the *very first stars in our Universe*, corresponds to the black→yellow transition in Fig. 1.
4. **IGM heating:** The IGM is heated, with the spin temperature now coupled to the gas temperature,  $T_K = T_S$ . Fluctuations are sourced by the gas temperature. As the gas temperature surpasses  $T_\gamma$ , the 21cm signal changes from absorption to emission, becoming insensitive to the actual value of  $T_S$  (see eq. 1). This epoch probes all processes which heat the IGM, *both astrophysical and cosmological*. The dominant source of heating is likely the X-rays from early accreting black holes (e.g. [7]) or from the hot interstellar medium (ISM; [9]). This epoch corresponds to the yellow→blue transition in the panels of Fig. 1.
5. **Reionization:** as the abundance of *early galaxies* increases, the IGM gradually becomes ionized, a process which is inside-out on large scales. Fluctuations during the advanced stages are dominated by the ionization field. The tomography of this process is sensitive to the nature and clustering of the dominant UV sources (e.g. [10]), as well as the evolution of inhomogeneous recombinations (e.g. [11]). The 21cm signal decreases, approaching zero.

This epoch corresponds to the blue→black transition in the panels of Fig. 1.

The first two stages (the Dark Ages) allow us to probe cosmology at redshifts much lower than recombination, while the last three stages are sensitive to early astrophysical sources (and sinks) of cosmic radiation fields. These last three stages will be observable with SKA1-LOW. Below, we focus on the astrophysical insight which can be gained from the CD and EoR (for cosmological insights, see §Cosmology from the EoR/CD). As a foreshadowing of how astrophysics can impact the CD and EoR signals, in the bottom panel of Fig. 1, we show an alternate model, in which the early galaxies formed later, but were much more efficient at producing hard X-ray photons, saturating the unresolved X-ray background (XRB; [3]) by  $z \approx 10$ ; the CD and EoR are dramatically different in these two models.

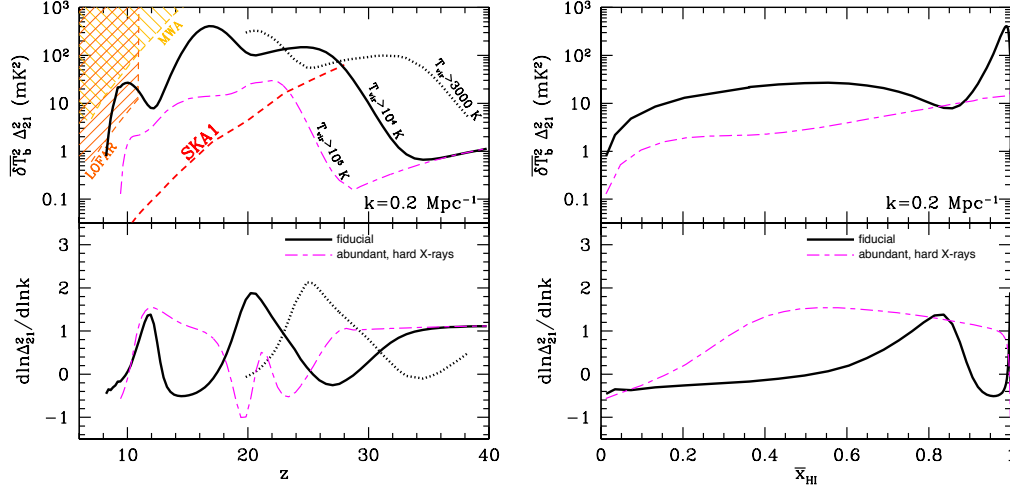
As our observable, we focus on the 21cm power spectrum. Although alternate statistics can provide complimentary observations, the power spectrum serves to quantify the main impact of astrophysics on the morphology of the 21cm signal. We define the 3D power spectrum as  $P_{21}(k, z) = k^3 / (2\pi^2 V) \delta\bar{T}_b(z)^2 \langle |\delta_{21}(\mathbf{k}, z)|^2 \rangle_k$ , where  $\delta_{21}(\mathbf{x}, z) \equiv \delta T_b(\mathbf{x}, z) / \delta\bar{T}_b(z) - 1$ . Moreover, when discussing detectability, we focus on the redshift evolution of large-scale power at  $k \approx 0.2 \text{ Mpc}^{-1}$  (e.g. [12]), roughly corresponding to the largest scales which should be relatively foreground free (e.g. [13]). Our models of the SKA1-LOW thermal noise are described in [14]. In brief, we take a fiducial 1000h observation, with a  $\Delta z = 0.5$  bandwidth and a frequency resolution of 1 kHz. 866 stations (each a 17x17 array of log-periodic dipoles) are placed using a Gaussian distribution with 75 % falling within 1000m of the center [15]. For most frequency bins and reasonable astrophysical parameters, the S/N can be dominated by cosmic variance (see [14] and Fig. 3). This suggest significant results already with a 50% early-science phase with SKA1-LOW, and also motivates a multi-tiered strategy combining several, moderately deep observations (as highlighted in the EoR/CD science goals; see also §7).

Unless stated otherwise, we quote all quantities in comoving units. The predictions below are consistent with recent Planck measurements of cosmological parameters [16].

Below we highlight how SKA1-LOW will allow us to study the galaxies and IGM during the CD/EoR. It will allow us to answer some of the most fundamental questions in astrophysical cosmology: *When did the first generations of galaxies appear? What were their UV and X-ray properties? What was the small-scale structure of the IGM?*

## 2. First, molecularly-cooled galaxies

The first galaxies likely formed at high redshifts,  $z > 30$  in very low mass halos  $M_{\text{halo}} = 10^{6-7} M_{\odot}$  (e.g. [17, 18, 19]), sometimes referred to as minihalos. Since the gas was primordial in composition, accretion and star-formation was governed by  $\text{H}_2$  cooling. Due to their shallow potential wells and inefficient cooling, the star formation inside minihalos is strongly susceptible to feedback effects, e.g.: (i) mechanical feedback from SNe explosions; (ii) X-ray heating; (iii) ionizing UV background (UVB); (iv)  $\text{H}_2$  dissociative radiation. The later, so-called Lyman-Werner background (LWB) is expected to eventually sterilize star-formation inside minihalos (e.g. [20, 21, 22]). In fact, even galaxies hosted by more massive, atomically-cooled halos (which are more resilient to feedback) could be sufficiently-abundant to establish a LWB strong enough (e.g. [23]) to sterilize star-formation inside minihalos by  $z \sim 20$  (e.g. [24, 25]; though see [26]).



**Figure 2:** Evolution of the 21cm power spectrum amplitude (*top panels*) and slope (*bottom panels*) at  $k = 0.2 \text{ Mpc}^{-1}$ . Evolution vs redshift ( $\bar{x}_{\text{HI}}$ ) is shown on the left (right). Solid black (dashed magenta) curves correspond to the fiducial (extreme X-ray) model shown in the top (bottom) panel of Fig. 1. The achievable thermal noise after a 1000h observation with SKA1-LOW is shown with the dashed red curve in the left panel; the analogous noise for the first generation instruments, LOFAR and MWA is marked by the checkered region in the top left corner of the panel. As discussed in the text, fiducial models show a three-peaked evolution in the large-scale 21cm power, corresponding to: (i) WF coupling, (ii) X-ray heating; and (iii) EoR. Models are taken from [1]. Also shown in the left panel is a model with an additional contribution at  $z \geq 20$  from  $\text{H}_2$  cooling halos with virial temperatures  $> 3000 \text{ K}$ .

Nevertheless, these fragile first galaxies are likely the ones which start the Cosmic Dawn. As mentioned above, this beginning is observable in 21cm through the WF coupling epoch, mostly driven by photons just redward of  $\text{Ly}\beta$  which emerge from galaxies and redshift into  $\text{Ly}\alpha$  resonance. These photons have mean free paths of  $\gtrsim 100 \text{ Mpc}$ , and can efficiently couple the spin temperature to the gas temperature, with 21cm being seen in absorption against the CMB. 21cm fluctuations during this epoch are driven by fluctuations in the strength of this coupling, i.e. the  $\text{Ly}\alpha$  background. Therefore, *the timing and duration of the initial rise and fall of the 21cm power tells us about the star-formation inside the very first, molecularly-cooled galaxies.*

This is highlighted by comparing the solid and dotted black curves in the left panels of Fig. 2. The solid black curve corresponds to a fiducial model, in which star formation only occurs in atomically-cooled galaxies. The dotted black curve assumes an additional contribution at  $z \geq 20$  from minihalos with virial temperatures  $> 3000 \text{ K}$ . The WF epoch starts earlier, and is more extended, driven by the slower evolution of the halo mass function on these scales. This delay is likely further extended by the feedback processes mentioned above. In this model, the peak in power associated with WF coupling would occur too early (at low frequencies) to be detectable with SKA1-low. However, the trough between the WF coupling and X-ray heating epochs should be detectable. Hence, if we do *not* detect a third peak in power at  $z \gtrsim 20$ , it would imply that minihalos are driving WF coupling.

Moreover, the *shape* of the power-spectrum at a given astrophysical milestone can tell us which halos hosted the first galaxies (e.g. [27]). As we shall also see below, radiation fields driven by more biased, rare galaxies result in less small-scale structure and associated 21cm power.



### 3. X-ray properties of early galaxies and IGM heating

As discussed above, X-rays from early galaxies are expected to heat the IGM to temperatures above the CMB well before the bulk of reionization. When the ionized fraction of the IGM surpasses a few percent, most of the X-ray energy gets deposited as heat through free-free interactions of the primary ionized electron (e.g. [28, 29, 30]). This makes X-rays more efficient as sources of heating than ionization. However, they could still significantly contribute to the EoR (in addition to the heating epoch), since the first, metal-poor galaxies are likely more X-ray luminous than local ones (per unit star formation rate; e.g. [31, 32]). Current limits do not rule out a strong redshift evolution in this efficiency of X-ray production (e.g. [33]). X-rays could also indirectly impact the EoR by raising the Jeans mass in the IGM. The resulting photo-heating feedback on sources might delay the EoR (e.g. [34, 1]), providing a window for a clean measurement of the matter power spectrum.

The interaction of X-rays with the IGM is characterized by a very large mean free path:

$$\lambda_X \approx 20 \bar{x}_{\text{HI}}^{-1} \left( \frac{E_X}{300 \text{eV}} \right)^{2.6} \left( \frac{1+z}{10} \right)^{-2} \text{cMpc}, \quad (3.1)$$

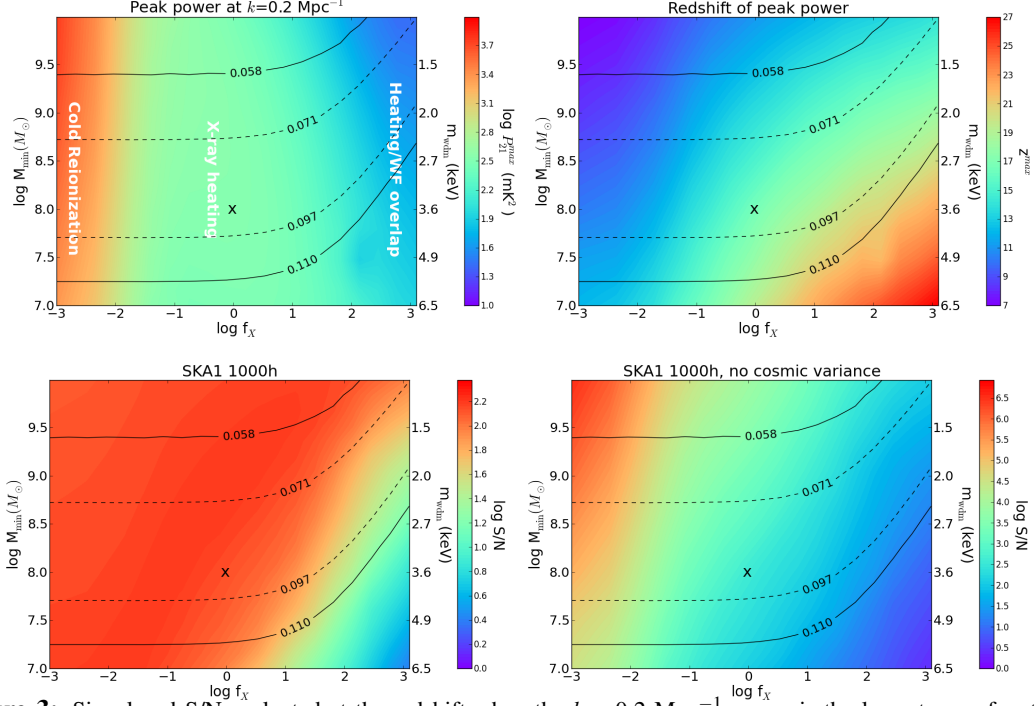
where  $\bar{x}_{\text{HI}}$  is the (volume) mean neutral fraction of the IGM and  $E_X$  is the photon energy. This means that only soft X-rays ( $E_X \lesssim \text{keV}$ ) interact with the IGM, making them relevant for 21cm studies. Without sub-keV X-rays, the EoR and IGM heating epochs would have dramatically different (smoother) morphologies (see the bottom panel of Fig. 1).

IGM heating during the CD is the most potent probe of early X-rays. X-rays from galaxies likely dominate the heating. First generation instruments will not detect this exciting epoch, for most reasonable scenarios [14]. The SKA1-LOW is poised to offer us the first glimpses into this exciting epoch, driven by the high-energy processes inside the first galaxies!

To further quantify this, in the top panels of Fig. 3, we show the peak amplitude and redshift of the peak amplitude of the large-scale 21cm power. We plot quantities in the parameter space of: (i) the minimum halo mass hosting star forming galaxies,  $M_{\text{min}}$ , and (ii) their X-ray efficiency,  $f_X$ . The fiducial value,  $f_X \equiv (N_X/0.25) = 1$  corresponds to  $N_X = 0.25$  X-ray photons per stellar baryon, consistent with empirical scaling relations from nearby star-forming galaxies (e.g. [35]). Similarly, the fiducial choice of  $M_{\text{min}} \sim 10^8 M_\odot$  corresponds to atomic-cooling galaxies (see §5 below). Increasing  $f_X$  shifts the X-ray heating epoch (and associated peak in power) towards higher redshifts, while increasing  $M_{\text{min}}$  shifts *all* astrophysical epochs towards lower redshifts (see top left panel of Fig. 2). Therefore, this 2D parameter space should span most of the variation in the signal.

Over a broad swath of parameter space, the large scale 21cm power during heating (driven by temperature fluctuations) peaks at a value of few hundred  $\text{mK}^2$  (c.f. [12]), an order of magnitude larger than the peak during the EoR (driven by ionization fluctuations). In the bottom panels of Fig. 3 we show the signal-to-noise (S/N) at which this signal can be detected with a 1000h observation of SKA1-LOW, with and without cosmic variance (left and right panels, respectively). *We see that the SKA1-LOW should easily detect X-ray heating throughout the parameter space.* Most scenarios are limited by cosmic variance, suggesting an observational strategy of sampling several independent fields, if higher S/N is desired, e.g. for imaging.

What are the likely sources of X-rays in the early galaxies? The bolometric X-ray luminosities of local, star-forming galaxies are dominated by bright, high mass X-ray binaries (HMXBs; e.g.

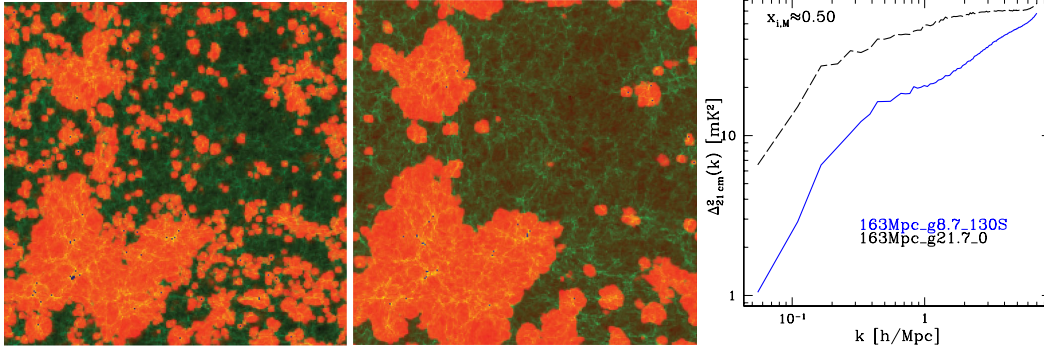


**Figure 3:** Signal and S/N evaluated at the redshift when the  $k = 0.2 \text{ Mpc}^{-1}$  power is the largest, as a function of the minimum halo mass hosting star forming galaxies,  $M_{\text{min}}$ , and their X-ray efficiency,  $f_X$ . *Top left:* peak signal, i.e. maximum amplitude of the 21cm power. *Top right:* redshift of the peak signal. *Bottom left:* corresponding S/N after 1000h with SKA1-LOW Phase 1. *Bottom right:* S/N neglecting the cosmic variance component:  $P_{21}/\sqrt{N_k}$ . A ‘fiducial’ model with  $M_{\text{min}} = 10^8 M_\odot$  and  $f_X = 1$ , is denoted with an ‘x’ in the panels. For the observing strategies used, see [14].

[36, 37]). However, the soft-bands relevant for X-ray heating,  $E_X \lesssim 2 \text{ keV}$ , have a comparable contribution from the hot interstellar medium (ISM; [38]). Because of the strong dependence of the X-ray mean free path on the photon energy (eq. 3.1), the X-ray spectral energy distributions (SEDs) of the first galaxies have a strong imprint on the 21cm fluctuations during the CD. In particular, scenarios in which X-ray heating is dominated by HMXBs should result in a factor of  $\sim 3$  less large-scale 21cm power than those dominated by the hot ISM [9]. This difference is easily identifiable with the SKA1-LOW, and importantly, is not degenerate with the galaxies’ X-ray luminosities. Hence, the SKA will be a powerful tool for studying the first galaxies and their high-energy processes.

#### 4. Radio loud sources at high-redshifts

Before the Universe was heated to temperatures greater than the CMB, the IGM could be visible in absorption along sightlines to high-redshift radio sources. This is the 21cm analogy of the well-studied Ly $\alpha$  forest. Although detecting features in the forest will be challenging even for SKA1-LOW (e.g. [39]), integration time can be reduced if one is only after a statistical measure, such as the increased variance along a sightline (e.g. [40]). In any case, detecting the 21cm forest along individual sightlines requires a radio-loud QSO at redshifts before heating (e.g. [39, 41]). It is not clear how likely this is, given that the population of bright QSOs declines rapidly towards high- $z$  (e.g. [42, 43]).



**Figure 4:** EoR morphologies driven by bursty star formation in atomically-cooled galaxies,  $M_{\min} \gtrsim 10^8 M_{\odot}$ , (*left panel*) compared to those driven by galaxies residing in  $M_{\min} \gtrsim 10^9 M_{\odot}$  (*middle panel*). Slices are 53 Mpc on a side, and correspond to the mid-point of reionization. The power spectra in the right panel are generated from 163 Mpc simulations with analogous source prescriptions. The power spectra corresponding to the two source models differ by factors of  $\sim 2$ – $3$  on large scales. Panels are from [46].

Chances can be improved however using the statistical imprint of a larger number of fainter QSOs. A high-redshift population of radio sources will introduce small-scale power in the 21cm power spectrum pre-heating. Temperature fluctuations driven by the X-ray emitting galaxies are expected to dominate the large-scale power, while the small-scale power ( $k \gtrsim 0.5 \text{ Mpc}^{-1}$ ) will allow us to constrain the population of radio loud active galactic nuclei at high redshifts [44].

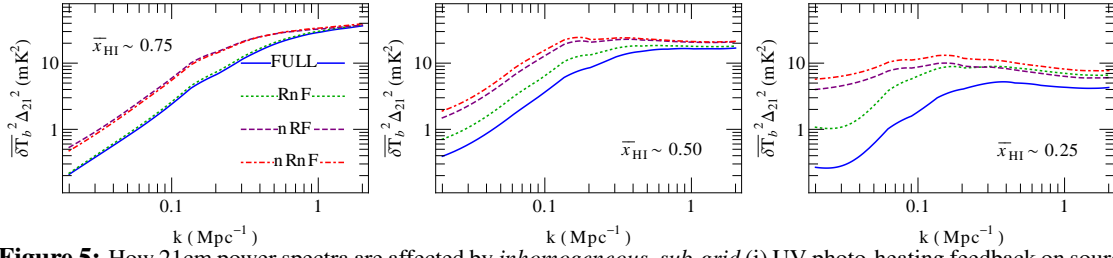
Dense structures such as damped Lyman alpha systems (DLAs), or even sterilized minihalos (mentioned above), would aid in the detection of the 21cm forest, if they contain enough cold, neutral gas. The resulting increased absorption features would be strongly imprinted in both individual sightlines (e.g. [45, 39]) as well as statistical detections [44], allowing us to additionally constrain the abundance of cold gas clumps at high- $z$ .

## 5. EoR sources

The SKA1-LOW we will map out the timing and duration of the EoR (see Fig. 2). The EoR is expected to be driven by UV photons from galaxies (e.g. [47]), and is characterized by some generic features in the evolution of 21cm power. Initially (following heating), there is a dramatic drop in large-scale amplitude and steepening of the slope during the early stages of reionization,  $\bar{x}_{\text{HI}} \gtrsim 0.8$ , when the burgeoning HII regions ‘cover-up’ the densest IGM patches (solid curve in the right panel of Fig. 2). This evolution can be understood (to first order, assuming  $T_S \gg T_{\gamma}$ ) as the transition of the 21cm power spectrum,  $P_{21} \approx P_{xx} - 2\bar{x}_{\text{HI}}P_{xd} + \bar{x}_{\text{HI}}^2 P_{dd}$ , from being dominated by the density power,  $P_{dd}$ , to the ionization power,  $P_{xx}$ ; the transition being governed by the negative contribution of the density-ionization cross spectrum,  $P_{xd}$  [48]. Subsequently, the fluctuations in the ionization field,  $P_{xx}$ , drive the large-scale 21cm power to a peak value during the mid-point of reionization.

However, this evolution is qualitatively different if hard ( $\gtrsim 1 \text{ keV}$ ) X-rays play a dominant role. The resulting weaker density-ionization cross-power and ionization power results in a much more gradual fall in power during the EoR, as well as a much lower overall amplitude (dashed curve in the right panel of Fig. 2). Both of these scenarios are easily identifiable with SKA1-LOW. More reasonable models, in which (soft) X-rays play a sub-dominant role in the EoR still show





**Figure 5:** How 21cm power spectra are affected by *inhomogeneous, sub-grid* (i) UV photo-heating feedback on sources and (ii) recombinations in sinks. The FULL (blue curves) model includes both effects via calibrated, sub-grid models. The RnF (green curves) model includes only recombinations. The nRF (purple curves) model includes only UVB feedback. The nRnF (red curves) model includes neither effects, and is comparable to current large-scale RT simulations of reionization (e.g. [51, 52]). Recombinations in sub-grid structures (and UVB photo-heating feedback to a lesser extent) can strongly suppress large-scale ionization structure (by factors of 2–5 throughout reionization). Figures are taken from [11].

a 21cm peak in large-scale power during the midpoint of reionization. However, the  $P_{xd}$  induced drop in power occurs earlier,  $\bar{x}_{\text{HI}} \sim 0.9$ , due to the pre-ionization from X-ray sources [1]. The 21cm probability distribution function (PDF) can also be used as a diagnostic, with X-rays decreasing the bi-modality of the PDF [12].

In addition to the timing and duration of the EoR, the 21cm signal can tell us about the nature of EoR galaxies. The EoR is likely driven by faint galaxies below the sensitivity limits of current and upcoming space telescopes, including *JWST*. The EoR galaxies formed through atomic cooling, requiring host halo virial temperatures of  $T_{\text{vir}} \gtrsim 10^4$  K ( $\gtrsim 10^8 M_{\odot}$  at  $z \sim 10$ ). However, the efficiency of star formation inside these dwarf galaxies at high- $z$  is highly uncertain. Moreover, it is likely that feedback processes (either through internal mechanical feedback, e.g. [49], or photo-heating feedback from the UVB; e.g. [50]), regulated the evolution of their star formation rate (SFR). The timing of the EoR signal, tells us when the dominant EoR sources appeared, which is a combination of their host halo masses and star formation efficiencies.

This degeneracy can be broken using the EoR morphology, which allows us to study the dominant host halo population of EoR galaxies. More massive host halos are more biased, with corresponding EoR morphologies with less small-scale ionization structure (see Fig. 4). This observation is however complicated by the fact that the bias of the halos evolves only weakly over this mass range [10], and by the fact that sinks of ionizing photons can also strongly impact the EoR morphology (see next section).

## 6. EoR sinks

By depleting the ionizing photon budget available to expand cosmic HII regions, recombining systems (largely comprised of so-called Lyman limit systems; LLSs) can have a large impact during (and following) cosmic reionization. Unfortunately, directly resolving such small ( $\lesssim \text{kpc}$ ) sinks of ionizing photons in large-scale reionization simulations is computationally impractical. Therefore, relatively untested sub-grid models must be used. SKA1-LOW will allow us to confront the predictions of such sub-grid models, offering insights into the small-scale baryonic physics of the IGM.

During reionization, gas is heated from low temperatures to  $\sim 10^4$  K. Depending on the level of pre-heating by X-rays, this might have a strong impact on the clumpiness of gas in the IGM (e.g. [53, 54]) and minihalos (e.g. [55, 56]). Subsequently, the gas establishes photo-ionization equilibrium with the local ionizing background, with most recombinations occurring in systems with densities large enough to partially self-shield (e.g. [57]). As a cosmological HII region grows, an increasing fraction of its ionizing photon budget is lost to balance recombinations, limiting its further growth [58]. This ‘slowdown’ is more dramatic for larger HII regions, since (i) they generally have a higher ionizing background, with ionization fronts driven into more dense, rapidly recombining systems; and (ii) their centers were the first to be reionized, thus having had enough time to recombine [11]. As large regions slow down, moderate HII regions can ‘catch’ up, further resulting in a more uniform size distribution even early in reionization.

As a result, recombinations strongly suppress the large scale ionization structure during the EoR. This can be seen from Fig. 5: comparing the green and red curves, we see that inhomogeneous recombinations can suppress the large-scale 21cm power by factors of 2–3. The effect is magnified by UVB feedback, which is also strongest inside large HII regions, whose cores ionized early enough for gas to respond to photo-heating. Comparing the blue and red curves, we see that the total impact of both (i) recombinations and (ii) UVB photo-heating feedback on sources results in a suppression of large-scale 21cm power by factors of 2–5. Neither effects are accounted for in most large-scale EoR simulations.

Although ‘directly’ detecting sinks through a corresponding rise in 21cm power on small-scales has been suggested [59], realistic models of their sizes and neutral gas mass make this direct detection unlikely even with the SKA1-LOW [11]. However, the indirect detection via the above-mentioned steep slope of the 21cm power spectrum should be easy to detect and can tell us about small-scale IGM structure.

## 7. Observing Strategy

Different observing strategies can impact the scientific return from SKA1-LOW. For example, the large-scale modes are dominated by sample (cosmic) variance, while the small-scale modes are affected by the intrinsic detector (thermal) noise. Increasing the total integration time of a single observed patch of the sky decreases the noise on small-scales, while instead observing many patches decreases the cosmic variance.

A quantitative analysis of EoR/CD constraints with different observing strategies has not been done yet. Qualitatively, we expect the largest constraining power to come from the redshift evolution of large-scale modes (e.g. [14, 60]). From figures 2 and 3 we see that a single 1000h observation with SKA1-LOW, even with a possible 50% reduction in area, will be limited by cosmic variance on large-scales. This suggests that optimal science returns would benefit from an increase in the total field of view, at the expense of a modest reduction in integration time. These considerations motivate SKA1-LOW’s current planned three-tiered observing strategy: (i) 1x1000h; (ii) 10x100h; (iii) 100x10h. The relative science gains from these strategies will depend on: (i) the effectiveness of foreground removal on large-scales; (ii) statistics used to characterize the cosmic signal; as well as (iii) the available instantaneous bandwidth probing the redshift evolution of the signal.

## 8. Conclusions

We summarize our main conclusions below.

- The timing and duration of the initial rise and fall of the 21cm power tell us about star-formation inside the very first, molecularly-cooled galaxies. Efficient star-formation inside these minihalos could source WF coupling fluctuations (driven by the Ly $\alpha$  background) as early as  $z \gtrsim 30$ . Hence if we do *not* detect a third peak in 21cm power on large scales at  $z \lesssim 30$  with SKA1-LOW, it would imply that minihalos are driving WF coupling.
- The peak *amplitude* of the large-scale 21cm power during X-ray heating is sensitive to the SEDs of the first galaxies (at the factor of few level), while the *redshift* of the peak tells us about their X-ray luminosity and host DM halos.
- The early stages of reionization,  $\bar{x}_{\text{HI}} \approx 0.8\text{--}0.9$ , are characterized by a drop in large-scale 21cm power amplitude, and a steepening of the power spectrum. The timing and duration of this feature tells us about the contribution of X-rays to the EoR.
- The midpoint of the EoR is characterized by a local maximum in large-scale 21cm fluctuations (driven by the ionization field), except if hard X-rays dominate reionization.
- The morphology of the EoR encodes information about the efficiency of star-formation inside galaxies, feedback processes and absorption systems.
- Absorption systems inhibit the growth of large HII regions, and could result in a dramatic reduction of large-scale 21cm power (by factors of 2–5). Therefore the steepness of the EoR power spectrum can tell us about the structure and evolution of these small-scale gas clumps.
- Radiation fields driven by more biased, rare galaxies result in less small-scale structure. Thus the 21cm power on small-scales encodes information on the host halos which host the dominant sources.
- The 21cm forest, if detected either through individual sightlines or statistically through the rise in 21cm power towards small-scales ( $k \gtrsim 0.5 \text{ Mpc}^{-1}$ ), will constrain the population of radio-loud AGN as well as cold gas clumps at high- $z$ .

## References

- [1] A. Mesinger, A. Ferrara, and D. S. Spiegel. Signatures of X-rays in the early Universe. *MNRAS*, 431:621–637, May 2013.
- [2] I. F. Mirabel, M. Dijkstra, P. Laurent, A. Loeb, and J. R. Pritchard. Stellar black holes at the dawn of the universe. *A&A*, 528:A149, April 2011.
- [3] R. C. Hickox and M. Markevitch. Resolving the Unresolved Cosmic X-Ray Background in the Chandra Deep Fields. *ApJ*, 661:L117–L121, June 2007.
- [4] S. R. Furlanetto, S. P. Oh, and F. H. Briggs. Cosmology at low frequencies: The 21 cm transition and the high-redshift Universe. *Physics Reports*, 433:181–301, October 2006.
- [5] S. A. Wouthuysen. On the excitation mechanism of the 21-cm (radio-frequency) interstellar hydrogen emission line. *AJ*, 57:31–32, 1952.
- [6] G. B. Field. . *Proceedings of the Institute of Radio Engineers*, 46:240, 1958.

- [7] S. R. Furlanetto. The global 21-centimeter background from high redshifts. *MNRAS*, 371:867–878, September 2006.
- [8] M. McQuinn and R. M. O’Leary. The Impact of the Supersonic Baryon-Dark Matter Velocity Difference on the  $z \sim 20$  21 cm Background. *ApJ*, 760:3, November 2012.
- [9] F. Pacucci, A. Mesinger, S. Mineo, and A. Ferrara. The X-ray spectra of the first galaxies: 21cm signatures. *ArXiv e-prints*, March 2014.
- [10] M. McQuinn, A. Lidz, O. Zahn, S. Dutta, L. Hernquist, and M. Zaldarriaga. The morphology of HII regions during reionization. *MNRAS*, 377:1043–1063, May 2007.
- [11] E. Sobacchi and A. Mesinger. Inhomogeneous recombinations during cosmic reionization. *MNRAS*, 440:1662–1673, May 2014.
- [12] S. Baek, B. Semelin, P. Di Matteo, Y. Revaz, and F. Combes. Reionization by UV or X-ray sources. *A&A*, 523:A4, November 2010.
- [13] J. C. Pober et al. Opening the 21 cm Epoch of Reionization Window: Measurements of Foreground Isolation with PAPER. *The Astrophysical Journal Letters*, 768(2):L36, 2013.
- [14] A. Mesinger, A. Ewall-Wice, and J. Hewitt. Reionization and beyond: detecting the peaks of the cosmological 21 cm signal. *MNRAS*; *arXiv:1310.0465*, February 2014.
- [15] P. Dewdney et al. SKA1 SYSTEM BASELINE DESIGN. *SKA1 SYSTEM BASELINE DESIGN* [http://www.skatelescope.org/wp-content/uploads/2013/03/SKA-TEL-SKO-DD-001-1\\_BaselineDesign1.pdf](http://www.skatelescope.org/wp-content/uploads/2013/03/SKA-TEL-SKO-DD-001-1_BaselineDesign1.pdf), pages 1–98, March 2013.
- [16] + Planck Collaboration. Planck 2013 results. XVI. Cosmological parameters. *ArXiv e-prints:1303.5076*, March 2013.
- [17] Z. Haiman, A. A. Thoul, and A. Loeb. Cosmological Formation of Low-Mass Objects. *ApJ*, 464:523–+, June 1996.
- [18] T. Abel, G. L. Bryan, and M. L. Norman. The Formation of the First Star in the Universe. *Science*, 295:93–98, January 2002.
- [19] V. Bromm, P. S. Coppi, and R. B. Larson. The Formation of the First Stars. I. The Primordial Star-forming Cloud. *ApJ*, 564:23–51, January 2002.
- [20] Z. Haiman, T. Abel, and M. J. Rees. The Radiative Feedback of the First Cosmological Objects. *ApJ*, 534:11–24, May 2000.
- [21] M. Ricotti, N. Y. Gnedin, and J. M. Shull. Feedback from Galaxy Formation: Production and Photodissociation of Primordial  $H_2$ . *ApJ*, 560:580–591, October 2001.
- [22] A. Mesinger, G. L. Bryan, and Z. Haiman. Ultraviolet Radiative Feedback on High-Redshift Protogalaxies. *ApJ*, 648:835–851, September 2006.
- [23] B. W. O’Shea and M. L. Norman. Population III Star Formation in a  $\Lambda$ CDM Universe. II. Effects of a Photodissociating Background. *ApJ*, 673:14–33, January 2008.
- [24] L. N. Holzbauer and S. R. Furlanetto. Fluctuations in the high-redshift Lyman-Werner and  $Ly\alpha$  radiation backgrounds. *MNRAS*, 419:718–731, January 2012.
- [25] A. Fialkov, R. Barkana, E. Visbal, D. Tseliakhovich, and C. M. Hirata. The 21-cm signature of the first stars during the Lyman-Werner feedback era. *MNRAS*, 432:2909–2916, July 2013.

- [26] K. Ahn, I. T. Iliev, P. R. Shapiro, G. Mellema, J. Koda, and Y. Mao. Detecting the Rise and Fall of the First Stars by Their Impact on Cosmic Reionization. *ApJ*, 756:L16, September 2012.
- [27] M. G. Santos, M. B. Silva, J. R. Pritchard, R. Cen, and A. Cooray. Probing the first galaxies with the Square Kilometer Array. *A&A*, 527:A93, March 2011.
- [28] J. M. Shull and M. E. van Steenberg. X-ray secondary heating and ionization in quasar emission-line clouds. *ApJ*, 298:268–274, November 1985.
- [29] S. R. Furlanetto and S. J. Stoevers. Secondary ionization and heating by fast electrons. *MNRAS*, 404:1869–1878, June 2010.
- [30] M. Valdés, C. Evoli, and A. Ferrara. Particle energy cascade in the intergalactic medium. *MNRAS*, 404:1569–1582, May 2010.
- [31] T. Fragos et al. X-Ray Binary Evolution Across Cosmic Time. *ApJ*, 764:41, February 2013.
- [32] A. R. Basu-Zych et al. The X-Ray Star Formation Story as Told by Lyman Break Galaxies in the 4 Ms CDF-S. *ApJ*, 762:45, January 2013.
- [33] M. Brorby, P. Kaaret, and A. Prestwich. X-ray binary formation in low-metallicity blue compact dwarf galaxies. *ArXiv e-prints:1404.3132*, April 2014.
- [34] M. Ricotti and J. P. Ostriker. X-ray pre-ionization powered by accretion on the first black holes - I. A model for the WMAP polarization measurement. *MNRAS*, 352:547–562, August 2004.
- [35] S. Mineo, M. Gilfanov, and R. Sunyaev. X-ray emission from star-forming galaxies - I. High-mass X-ray binaries. *MNRAS*, 419:2095–2115, January 2012.
- [36] M. Gilfanov, H.-J. Grimm, and R. Sunyaev.  $L_X$ -SFR relation in star-forming galaxies. *MNRAS*, 347:L57–L60, January 2004.
- [37] S. Mineo, M. Gilfanov, and R. Sunyaev. X-ray emission from star-forming galaxies - I. High-mass X-ray binaries. *MNRAS*, 419:2095–2115, January 2012.
- [38] S. Mineo, M. Gilfanov, and R. Sunyaev. X-ray emission from star-forming galaxies - II. Hot interstellar medium. *MNRAS*, 426:1870–1883, November 2012.
- [39] K. J. Mack and J. S. B. Wyithe. Detecting the redshifted 21 cm forest during reionization. *MNRAS*, 425:2988–3001, October 2012.
- [40] C. L. Carilli, N. Gnedin, S. Furlanetto, and F. Owen. Observations of HI 21-cm absorption by the neutral IGM during the epoch of re-ionization with the Square Kilometer Array. *NAR*, 48:1053–1061, December 2004.
- [41] B. Ciardi et al. Prospects for detecting the 21 cm forest from the diffuse intergalactic medium with LOFAR. *MNRAS*, 428:1755–1765, January 2013.
- [42] C. J. Willott et al. Four quasars above redshift 6 discovered by the Canada-France High-z Quasar Survey. *ArXiv e-prints 0706.0914*, June 2007.
- [43] R. J. Wilman et al. A semi-empirical simulation of the extragalactic radio continuum sky for next generation radio telescopes. *MNRAS*, 388:1335–1348, August 2008.
- [44] A. Ewall-Wice, J. S. Dillon, A. Mesinger, and J. Hewitt. Detecting the 21cm Forest in the 21 cm Power Spectrum. *ArXiv e-prints:1310.7936*, October 2013.
- [45] S. R. Furlanetto and A. Loeb. The 21 Centimeter Forest: Radio Absorption Spectra as Probes of Minihalos before Reionization. *ApJ*, 579:1–9, November 2002.



- [46] I. T. Iliev, G. Mellema, P. R. Shapiro, U.-L. Pen, Y. Mao, J. Koda, and K. Ahn. Can 21-cm observations discriminate between high-mass and low-mass galaxies as reionization sources? *MNRAS*, 423:2222–2253, July 2012.
- [47] M. McQuinn. Constraints on X-ray emissions from the reionization era. *MNRAS*, 426:1349–1360, October 2012.
- [48] A. Lidz, O. Zahn, M. McQuinn, M. Zaldarriaga, and L. Hernquist. Detecting the Rise and Fall of 21 cm Fluctuations with the Murchison Widefield Array. *ApJ*, 680:962–974, June 2008.
- [49] V. Springel and L. Hernquist. The history of star formation in a  $\Lambda$  cold dark matter universe. *MNRAS*, 339:312–334, February 2003.
- [50] A. A. Thoul and D. H. Weinberg. Hydrodynamic Simulations of Galaxy Formation. II. Photoionization and the Formation of Low-Mass Galaxies. *ApJ*, 465:608–+, July 1996.
- [51] O. Zahn, A. Mesinger, M. McQuinn, H. Trac, R. Cen, and L. E. Hernquist. Comparison of reionization models: radiative transfer simulations and approximate, seminumeric models. *MNRAS*, 414:727–738, June 2011.
- [52] S. Majumdar, G. Mellema, K. K. Datta, H. Jensen, T. R. Choudhury, S. Bharadwaj, and M. M. Friedrich. On the use of semi-numerical simulations in predicting the 21-cm signal from the epoch of reionization. *ArXiv e-prints:1403.0941*, March 2014.
- [53] A. H. Pawlik, J. Schaye, and E. van Scherpenzeel. Keeping the Universe ionized: photoheating and the clumping factor of the high-redshift intergalactic medium. *MNRAS*, 394:1812–1824, April 2009.
- [54] J. D. Emberson, R. M. Thomas, and M. A. Alvarez. The Opacity of the Intergalactic Medium during Reionization: Resolving Small-scale Structure. *ApJ*, 763:146, February 2013.
- [55] P. R. Shapiro, I. T. Iliev, and A. C. Raga. Photoevaporation of cosmological minihaloes during reionization. *MNRAS*, 348:753–782, March 2004.
- [56] I. T. Iliev, P. R. Shapiro, and A. C. Raga. Minihalo photoevaporation during cosmic reionization: evaporation times and photon consumption rates. *MNRAS*, 361:405–414, August 2005.
- [57] M. McQuinn, S. P. Oh, and C.-A. Faucher-Giguère. On Lyman-limit Systems and the Evolution of the Intergalactic Ionizing Background. *ApJ*, 743:82, December 2011.
- [58] S. R. Furlanetto and S. P. Oh. Taxing the rich: recombinations and bubble growth during reionization. *MNRAS*, 363:1031–1048, November 2005.
- [59] T. R. Choudhury, M. G. Haehnelt, and J. Regan. Inside-out or outside-in: the topology of reionization in the photon-starved regime suggested by  $\text{Ly}\alpha$  forest data. *MNRAS*, 394:960–977, April 2009.
- [60] J. C. Pober and other. What Next-generation 21 cm Power Spectrum Measurements can Teach us About the Epoch of Reionization. *ApJ*, 782:66, February 2014.